It's About Time: Matching English Learners and the Ways they take Tests by Using an Online Tool to Properly Address Individual Needs

Rebecca J. Kopriva, Therese Gleason Carr, and Julia Lara

Need for Improved Accommodations Decision-Making for English Learners

Since the passage of *No Child Left Behind* (2001), progress has been made in both research and practice in properly identifying English learners and meaningfully including them on large-scale assessments for accountability purposes, yet many challenges and questions remain. The nature of the decision-making process required to appropriately match individual students to particular accommodations is complex, given the many and varied background characteristics of English learners as well as the limited–yet growing–research-based guidance for educators who may not have had the opportunity to receive adequate training in assessment and the assignment of accommodations. In particular, the highly diverse nature of the ELL population–with students from a multitude of linguistic, cultural, academic, and testing backgrounds– requires a more systematic means of offering appropriate accommodations than the methods typically used in schools and in large-scale assessments. Further, it is difficult to select appropriate testing accommodations when information about the usefulness of new and existing accommodations is still emerging. Although various accommodations have been adopted across states to include ELLs in standardized testing, results about their usefulness are still mixed, particularly if students are not matched appropriately to accommodations based on their specific background and needs. Thus, there are currently a number of necessary and worthy research agendas spanning both past, present, and future issues within the field of valid testing accommodations and decision-making for English learners.

First, additional research into the validity and efficacy of the accommodations that are currently widely available and in use across most states is needed. There is also general agreement among measurement and ESOL researchers that the current domain of testing accommodations for English learners, which was initially based largely on the existing more well-established paradigm for students with disabilities, lacks sufficient accommodations offerings that adequately address the common denominator of an otherwise highly heterogeneous population of students: linguistic and cultural challenges. Similarly, there is consensus that English learners encounter unique barriers to accessing assessments and demonstrating what they know and can do in regard to academic content standards, challenges that are directly related to their simultaneous charges of acquiring proficiency in the English language and learning academic content. Consequently, experts also agree that innovative accommodations that more effectively address the unique linguistic and cultural needs of English learners are clearly warranted.

Such innovations often lend themselves naturally to computer-based and/or adaptive platforms, which allow maximum flexibility compared to the static paper-based instruments still in use by a majority of state and local education agencies. On a positive note, a number of innovative new tests and tools exploiting the benefits afforded by computer technology are currently in development. However, while these new tools are increasingly available, research empirically demonstrating their validity, not to mention their availability and state and local stakeholders' capacity to use them, is still nascent. Further, there is less of a clear idea as to how best to match appropriate accommodations and/or other testing strategies, both those that are existing/widely available and more emergent approaches, to individual English learners. Numerous researchers have demonstrated not only the lack of consistency in teacher accommodations decision-making but also the consequences of providing learners with inappropriate, invalid, and/or unnecessary accommodations—and most educators are all too familiar

with the myriad potentially negative repercussions, including long-term impacts, that result from inaccurate test scores. Unfortunately, the most well-intentioned, defensible and even advanced approach is unlikely to succeed if it is ill applied to a particular circumstance or offered to a student for whom it is not appropriate. As Kopriva (2008) puts it, "even as tests may be developed and accommodated to specifically address the needs of English language learners (ELLs), if there is no technically rigorous mechanism in place to get the specific methods to the specific students who need them, it is argued that these efforts have little effect" (p. number from book?). In other words, it is essential to consider how best to identify, assign, and combine particular accommodations based on specific individual background characteristics of English learners if the desired effect of reducing systematic error on assessments for English learners—and all the access, fairness, and equity issues this goal encompasses—is to be achieved.

Thus, the focus of this paper will be to explore the most relevant and important student characteristics among English learners of different proficiency levels and how they interact with various aspects of current accommodations in use today in such a way as to yield valid and effective accommodations packages for individual learners. Distinguishing features among and within English learners with different proficiency levels will be discussed as they are relevant to accommodations decision-making. "Active" or especially relevant aspects of particular accommodations that define which skills are needed to access the accommodations themselves will be described and differentially linked to English learners in various proficiency level groups. Attention will also be devoted to considering how sets of accommodations, and in particular their active characteristics, work together to provide the most effective avenue for particular students to access large-scale assessments. To illustrate these principles, several cases will be presented depicting profiles of English learners with varying proficiency levels and other relevant background characteristics that should condition valid accommodations decision-making; a discussion modeling the type of decision-making required to appropriately match English learners to packages of accommodations will be included for each case. Finally, the Selection Taxonomy for English Language Learner Accommodations (STELLA), an online tool designed to assist educators in differentially selecting appropriate accommodations for individual English learners, will be presented as one approach to thoughtfully and meaningfully making accommodations decisions for English learners.

Need to Assign Accommodations Differentially for English Learners

In the large-scale assessment literature, research on the most effective accommodations for both English learners and students with disabilities is ongoing. Concurrently, movement is also being made to appropriately integrate these students in standardized testing systems designed for the broad range of students being educated in U.S. schools (No Child Left Behind, 2001). It is generally believed that accommodations should be selected to meet the needs of the individual student. However, research confirms that one cannot validly assign accommodations to groups of students based on some broad classification or status (Sireci, Li, & Scarpati, 2003). Emerging work suggests that systematic methods of assignment may work better than relying on current policy approaches. Further, it presents evidence that using systematic methods to match the particular needs and strengths of individual students to specific accommodations which address these needs and strengths may increase validity and be superior to using educator directed decision-making alone (Koran, Kopriva, Emick, Monroe & Garavaglia, 2006; Kopriva, Emick, Hipolito-Delgado, & Cameron, in press; Tindal, 2006). (Kopriva, 2008)

Overview of Key Relevant Student Factors for Accommodations Decision-Making

Given the heterogeneity of English learners as well as the complexity of the process of second language acquisition—particularly when undertaken while simultaneously learning academic content—there are a literally a multitude of student background characteristics and factors that are likely relevant to large-

scale testing for individual learners. In order to ensure that any accommodations decision-making approaches presented to educators and policy makers are feasible in terms of time and resources, however, it is necessary to cull from this large list the most critical student factors to consider. A number of researchers have narrowed the list of background characteristics by focusing specifically on those that are most relevant to accommodations selection. A brief survey of the literature indicates a high level of agreement as to which student factors are—or should be--most salient to accommodations decision-making.

For example, Butler and Stevens (1997) identified the following three background factors that they believe to be most salient in accommodation selection: English language proficiency, prior formal schooling, and length of time in the U.S. Similarly, Abedi (2007) and Rivera and Collum (2006) mostly emphasize English language proficiency levels and the language of instruction. Kopriva and others conducted an extensive review of the literature as well as qualitative research activities including cognitive labs, interviews and focus groups; this group of researchers highlighted several key student background characteristics that appear to be most salient for matching ELLs to appropriate accommodations. The relevant variables from these studies include levels of proficiency in both English and the home language or L1 (in all four domains of reading, writing, listening, and speaking) and a variable called cultural proximity, or the degree to which the student's schooling and testing experiences in the home country are similar to U.S. schooling and testing systems. These authors also underscore the importance of teacher judgments of student needs and strengths as well as ongoing testing practices and formats that appear to benefit the student (See Kopriva, Hedgspeth, Koran, & Carr, 2007, Kopriva et al., 2005a; and Winter et al., in press). Kopriva et al. and Winter are not the only researchers to examine these types of factors; others, including including Saville-Troike (1991) and Butler and Stevens (1997), have also isolated these types of factors as they relate to accommodations decision-making.

Cultural proximity, as Kopriva and others have defined it, is an aggregate variable which refers to the similarity between the student's native country schooling (if applicable) and selected home school-like experiences, relative to U.S. schooling and testing opportunities. Aspects of this variable are highlighted by several researchers, including Saville-Troike (1991) and Butler and Stevens (1997). The type of information collected by Kopriva et al. focuses on native country schooling experiences, including resources, time as it relates to the academic year/amount of time spent in class per day, pedagogical approaches, and purposes of testing. Kopriva (2008) has also highlighted the importance of collecting and analyzing information about not only the amount of time ELLs have spent in U.S. schools but also the consistency of the students' attendance and experiences, given that gaps in schooling related to agricultural and other seasonal work cycles are a relevant issue for many students in this population.

English Language Proficiency as a Relevant Student Factor

Identifying and classifying English learners into the appropriate English language proficiency group is a tricky prospect, even with the many new and improved tools and tests on the market to assist educators in this task. The process is complex because of the diversity of English learners in general, the often asynchronous nature of second language acquisition across domains (reading, speaking, writing, listening), and because English language proficiency is a "moving" target: individual learners may proceed at different paces from each other as well as more quickly/slowly during different phases of their own language development. Furthermore, English language proficiency may mean different things to different people; in the past, for example, social, functional speaking skills were emphasized in teaching ESOL rather than academic language. Increasingly, educators are now more aware of the distinction between Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP) and many language placement and assessment instruments now reflect this.

Nevertheless, care must be taken to critically read and understand the varying classifications of ELP and their purposes as used by different groups in order to ensure their appropriate application.

For instance, many tests report composite ELP scores for students as well as composite domains, e.g., "comprehension", which includes listening and ???. Educators need to be aware of not only how test and/or other data are aggregated for these purposes but also how to interpret this information—and whether, based on their own first-hand knowledge of individual learners, such information is valid. Test scores or other "objective" measures of ELP should not be considered the "end all, be all" but rather should be critically considered by educators in light of their own expertise and knowledge of what individual students know and can do in content areas as well as where they are in the process of developing English. In order to empower teachers and other educators to do this effectively, it is helpful to first consider what are, generally, the most essential background characteristics of ELLs in different proficiency categories as they relate to accommodations decision-making. After such an understanding is gained, educator may feel more confident formulating and incorporating their own judgments in thinking about how to best assign accommodations for individual English learners.

L1 Proficiency as a Relevant Student Factor

L1 levels reflect the student's ability to effectively use his or her native language relative to students at the same grade level studying in full-time academic programs in the student's native country (even if this student has never attended school in the native country). It may be helpful to think about L1 proficiency in terms of four levels: 1) ***Well below grade level***, which means that the student has little or no proficiency in his or her native language or has proficiency substantially below the typical level of L1 proficiency observed at the same grade level in a full- time academic program in the native country; 2) ***Below grade level***, which indicates that the student's L1 proficiency is somewhat below what one would expect of same grade-level peers enrolled in full-time school in the native country; 3) ***At grade level***, which indicates that the student can communicate in L1 at a level that is competitive in meeting grade-level expectations in the native country; and 4) ***Above grade level***, which indicates that this student demonstrates an advanced level of proficiency in his or her native language as compared to his or her peers in the native country.

While students may be "limited" in English proficiency, if some English proficiency can be assumed (usually reading), bilingual children often have greater overall linguistic expertise and advantage compared to their monolingual peers who are having trouble reading tests (Kester & Peña, 2002). Unequal proficiency in native language abilities may impact testing and necessitate accommodations to compensate for this strain. Research in this area suggests that using native language proficiency represents a collaborative "meaning-making" process with which to facilitate the measurement of academic content (Knapp, 1995; Ruddell, 1995). Oral L1 proficiency is particularly important if they are being instructed in L1 or if they are not yet literate enough in their native language. Additionally, oral L1 aid for non-glossed words may be useful for students with no to little written reading proficiency in their native language and if they have been learning content in English for a short period of time (Mann, Emick, & Kopriva, 2006).

If and as they have some degree of written native language proficiency, written aid in L1 can supplement tests in English for those in English-speaking content classrooms. This because they can get clues about English item meaning from L1 cognates and other linguistic components (August, 2004). Several studies (for example Collier & Thomas, 2002) have shown that as students are taught literacy skills in their home

language this seems to have an effect on educational outcomes such as test scores, academic performance, etc. Even when ELL students are rather literate in English, Solano-Flores and others (2002) suggests that native language aids can continue to be helpful. Finally, several authors (including Rivera & Stansfield, 2001; Duncan et al., 2005) emphasize that native language and dual language forms, as well as use of language aids (particularly if they are written), takes more time than if students are proficient in English and using English forms. For instance, forms in Spanish are approximately one-third longer than their counterparts in English. Knowledge about these components in ELLs' home language reveals details about their level of proficiency which offers relevant information about these students' specific test-taking needs (LaCelle-Peterson & Rivera, 1994; Durán, 1997; Butler et al., 2001).

Unequal proficiency in native language abilities may impact testing and necessitate accommodations to compensate for this strain. Research in this area suggests that using native language proficiency represents a collaborative "meaning-making" process with which to facilitate the measurement of academic content (Knapp, 1995; Ruddell, 1995).

Oral L1 proficiency is particularly important if they are being instructed in L1 or if they are not yet literate enough in their native language. Additionally, oral L1 aid for non-glossed words may be useful for students with no to little written reading proficiency in their native language and if they have been learning content in English for a short period of time (Mann, Emick, & Kopriva, 2006).

While students may be "limited" in English proficiency, if some English proficiency can be assumed (usually reading), bilingual children often have greater overall linguistic expertise and advantage compared to their monolingual peers who are having trouble reading tests (Kester & Peña, 2002).

Cultural Proximity as a Relevant Student Factor

Cultural proximity, as defined by Kopriva et al., is a measure of how similar the student's home country schooling system, including academic assessment, was to the U.S. school system. There are three levels for cultural proximity: High (very similar to U.S. culture), Medium (somewhat similar) and Low (not similar or very dissimilar). It has been argued that to better understand children's learning, one must first appreciate the cultural environment in which such learning occurs (Boykin, 1995; Greeno, 1989; Rogoff, 1990). Research indicates that students whose home culture is similar to their "school culture" (typically based on mainstream, middle-class U.S. culture, according to Farr and Turnbull (1997)) tend to fare better educationally than those whose home and school cultures differ substantially (Boykin and Bailey, 2000; Grenfell and James, 1998). Clearly, language minority students come from many different cultural backgrounds and life experiences that affect their educational needs (e.g. Hakuta & Beatty, 2003; LaCelle-Peterson & Rivera, 1994).While many ELLs share the common goal of improving their English language proficiency, they differ in cultural background, home language, family and educational history and so on. ELLs are from various distinct communities, and research in this area asserts that each of these communities may interact with schooling very differently (Au, 1980; Au & Jordan, 1981; Butler & Stevens, 1997; Jacob & Jordan, 1987; LaCelle and Rivera, 1994; Ogbu, 1987; Phillips, 1983; Trueba, 1989). Therefore, several of these researchers suggest it is important that information be obtained about some of the cultural experiences of the student, especially as they relate to their schooling experiences that are key to understanding classroom and large scale testing environments. Outcomes from this body of research also find that careful consideration should be paid to particular components of the experiences of the students' parents and the expectations of their home communities to the extent that these are different than the mainstream U.S. testing perceptions. Further, knowledge of student strengths and challenges given their home communities and cultural backgrounds seems to be

salient. Malagon-Sprehn et al. (2004) reported that that, because a large percentage of ELLs were native born or in home communities somewhat isolated from mainstream U.S. culture, the culture of their home communities seems to have an impact in some of the same way that experiences in home countries might and they suggest that this as well as home country experiences should be considered. Butler and Stevens (1997) report that length of time in the U.S. as well as amount and type of prior formal schooling is a necessary but not sufficient "proxy" for the degree of cultural connections and disconnections a student has with the U.S. schooling and testing culture. Additionally, Lara and August (1996) maintain that consistency in U.S. schooling is closely related and should be considered. The time and school consistency in U.S. are often used to suggest the degree of English language proficiency as well.

Time in U.S. as a Relevant Student Factor

In considering the relevance of the amount of time an ELL student has spent in U.S. schools as this relates to accommodations decision-making, it is important for educators to not just consider the amount of time but also the consistency and quality of the time spent enrolled in U.S. schools.

The amount of time a student has been in the U.S. reflects how much opportunity the student has had to adapt to U.S. culture. When thinking about time as it relates to accommodations decision-making, it is helpful for educators to take into account how long the child has been in the United States, as well as his/her time spent in U.S. school systems. In general, the more time the student has spent in U.S. schools, the more likely the he or she is to be familiar with U.S. schools and assessment practices. For example, if a student has been in the U.S. for more than three years and has missed less than two months of school for less than 2/3 of the year, the student is likely to have a high level of familiarity with U.S. schooling and testing practices. By contrast, a student who has been enrolled in U.S. schools between two and three years but has missed more than two months of school per year at least two of the years would almost certainly have significant gaps not only in opportunity to learn academic content but also would have had less opportunity to become acculturated and comfortable with the U.S. school system, including testing purposes, items, and formats.

U.S. Schooling Experience as a Student Factor

When making accommodations decisions, it is important to consider the degree of exposure that an individual student has had to testing protocol in the United States. Assessment may differ in procedures from country to country, and depending on the circumstances a student has encountered thus far in his or her schooling experience, the student may be quite familiar with the types of assessment procedures used in the U.S., or he or she may have little experience with them. Appropriately matching English learners to accommodations packages requires considering whether students are very familiar with test procedures common in the U.S., somewhat familiar with some U.S. test procedures, or not very familiar/completely unfamiliar with U.S. test procedures—and thinking about how the student's level of testing experience impacts his or her need for accommodations in order to demonstrate knowledge and skills on state content assessments. Similarly, teachers should think about what particular needs the student has in the classroom, as well as any compensatory strengths he or she possesses. This information should also be weighed as it relates to experience with U.S. testing and interacts with the other factors discussed here.

Overview of Promising Accommodations

The following accommodations have been identified in the research literature as the most promising for English learners. Based on research findings and practitioner experiences, Rivera and Collum (2006) and Abedi (2007) have recently presented the subset of accommodations that seem to be the most salient for English learners. Pennock-Roman and Rivera (2006) have also reported effect size benefits in their

meta-analyses of relevant research to date. As noted throughout, the research findings are at various levels of consensus on the most beneficial accommodations for this population. However, this is largely because the field has continued to learn about how to evaluate effectiveness in general (Sirici, Li, & Scarpati, 2003), and particularly how to determine effectiveness for students with specific needs.

Promising Accommodations for English Language Learners

- Forms
  - Access-based form in English
  - Native language or dual language forms as available
- Tools
  - Bilingual word list, general or test specific
  - Picture-word dictionary
  - Problem solving tools
- Administration
  - Oral English
  - Oral home language
  - Small group
  - Language liaison
  - Extra time
  - More frequent breaks
- Response
  - Written in native language or code switching
  - Oral English
  - Oral in native language or code switching
  - Demonstrated or modeled response

Active Characteristics of Accommodations

Active characteristics are those aspects of the accommodations which define which skills are needed to access them. For instance, oral administration in English requires that a student have sufficient auditory capability. Further, they should be able to understand the language of English at the level of language complexity in the written text which is being recited.

Kopriva (2005a) reported that a task analysis of the active characteristics was completed for 15 promising accommodations. Active characteristics are those aspects of the accommodations which define which skills are needed to access them. For instance, oral administration in English requires that a student have sufficient auditory capability. Further, they should be able to understand the language of English at the level of language complexity in the written text which is being recited. The 15 accommodations were selected based on literature reviews and pilot findings in this project and another related study (Kopriva & Mislevy, 2005). For the most part, they were consistent with what was cited by Rivera and Collum and Abedi. Once the analyses of the accommodations were completed, the options were preloaded into the assignment system which was being built. The list of the accommodations in Table 10.1 has been adapted from this work. The system will be discussed later in this chapter, but the logic essentially begins with a screen which evaluates the student's capability of taking the assessment under standard testing conditions. These conditions include a standard form as well as the testing procedures typically used in conducting large-scale testing. Therefore, accommodations are activated in the system when minimum thresholds for taking the assessment under standard conditions are not

meant. Because the accommodations are discussed at length in other chapters, a discussion of the options, their research base, and their effectiveness will not be repeated here.

Understanding How Accommodations Packages Work Together

Studying test accommodations and understanding how they work can be quite complicated. Tindal and Ketterlin-Geller (2004) note that in accommodations research, "few accommodations can be viewed as single, isolated variables," and that "…accommodations are best thought of as a package in which no one accommodation is ever studied very well in isolation" (p.7). On one hand, test changes that we think of as a single accommodation, such as reading a test aloud to a student, may have multiple facets that remove or contribute construct irrelevant variance for a particular student. For example, if a test is read aloud to a group of students, it is possible that students may benefit from the effect of having the test administrator pace them through the test, which is different from the benefit derived by having the reading demand of the test eliminated or reduced (Hollenbeck, Rozek-Tedesco, Tindal, & Glasgow, 2000; Weston, 2003).

Furthermore, groups of such accommodations may also be thought of as packages (Hollenbeck, Rozek-Tedesco, Tindal, and Glasgow, 2000), as the multiple facets of various accommodations interact to produce a net increase or decrease in construct irrelevant variance. Linguistic accommodations in particular (such as different forms of the test that incorporate linguistically-oriented changes in the presentation of items or the availability of language aids, for instance a bilingual word list or decreased language complexity) function most optimally as packages. For example, side-by-side (dual language) administrations require extra time because working back and forth between the two language versions of each item takes more time than completing a comparable test in one language (Choi & McCall, 2002). As Chapter 1 discusses, this seems to be particularly true for ELL students with little English proficiency who may be especially affected by their previous schooling environments, past experiences, and current home environments which may be very different than those associated with mainstream U.S. students.

One may think of multiple possible combinations of accommodations that function together to meet the student's need. One package of accommodations may be considered the preferred package for a particular student although other packages compensate for the student's need nearly as well. For example, an ELL may have a level of English proficiency such that a bilingual word list may not be needed if the student is to receive a form of the test where items are presented in plain or simplified English or in their native language. However, if that student were to receive a standard English version of the test form, he or she would need the bilingual word list to help him or her access the items. In selecting accommodations for ELLs most accommodation researchers are now recognizing that the package must be tailored to the needs of the individual student. Further, over-accommodating can be just as problematic for the student as not providing the needed accommodations at all. Administering a test with all possible accommodations may be overwhelming and possibly counterproductive (Kopriva, 2005a) Administering a test with improper "bells and whistles" (particularly un-necessary additions to the test items or forms) that are not needed can be distracting (for instance, see Sharrocks-Taylor & Hargreaves, 1999).

Overview of Distinguishing Features of Low, Medium, and High ELP Groups and How These Features Interact with Accommodations

For purposes of this paper, since there is such variation in the field on the classification of ELLs, English language proficiency first will be discussed broadly in terms of low, medium, and high English language proficiency groups and some of the most salient distinguishing characteristics of students in these groups. Then, special attention will be paid to the characteristics of the "in between" groups of learners, or those who fall somewhere between low and medium ELP, and medium and high ELP. These "in-

between" groups that can be particularly challenging to classify include those that contain students who have neither low nor medium ELP as defined above, and those students who have neither medium or high ELP as defined above. It is important to be aware of the types of characteristics these two liminal groups possess so that these students do not fall through the proverbial cracks when accommodations and other instructional decisions are made. A contributing factor to confusion about these groups' actual ELP levels—and thus how best to serve their needs—may hinge in part on sometimes misleading nomenclature used to refer to these students. While these descriptions are by no means comprehensive, they are intended as a summary of the most important and relevant background characteristics to consider when thinking about matching the best accommodations package to English learners with varying degrees of English proficiency.

An additional objective of reviewing the types of general characteristics that students in the same ELP groups generally share relative to their ELP development is to provide a foundation from which to consider other factors that distinguish not just between proficiency groups but also among students within these groups. It is necessary to understand the characteristics of English learners in different proficiency groups in order to begin thinking about how to match accommodations appropriately to these students, both between and within groups. Below, "between-group" considerations will be discussed first, in terms of matching categories of ELLs with certain distinguishing features with the categories or types of accommodations that are likely most salient for them. The following section of this paper will address "within-group" accommodations selection, or matching specific packages of accommodations with individual English learners with particular background characteristics, needs, and strengths.

Understanding How Accommodations Packages "Match" with Different ELP Proficiency Groups

In order to move toward a more effective and thoughtful decision-making process for accommodations selection for ELLs, it is helpful to begin to consider the extent to which certain types of accommodations and packages of accommodations are, broadly speaking, more appropriate for ELLs at different ELP levels. One of the goals of this paper is to empower educators and policymakers to critically think about accommodations selection in order to be able to effectively advocate for the accommodations packages that are truly customized for their students on an individual level. This approach is diametrically opposed to the frequently employed, well-intentioned, yet inappropriate and often deleterious "blanket" or "cafeteria" approach of giving ELLs with highly divergent background characteristics as many accommodations as are currently allowed. Given this stated goal of education and advocacy for educators and policymakers, recommended accommodations packages for groups of students and for individual ELLs with particular profiles will be presented mainly in terms of what research thus far has indicated are best-practice—and thus "should" be essential or required—accommodations. Some brief discussion of possible "second-choice" alternate accommodations packages that may be more widely available/allowable will follow, but the emphasis will be mainly on the recommended, empirically based packages most likely to yield more valid inferences about student knowledge and skills in content areas.

Distinguishing Features of ELLs with *Low* English Language Proficiency

English learners in the low proficiency group range from having no English to demonstrating a very minimal understanding and use of English. In practice, the low ELP group typically encompasses students who are newly arrived in the U.S. or those students who have been enrolled in U.S. schools (and have had no previous English instruction) for less than or equal to one year. These students are likely to test into the lowest level of English proficiency on ELP placement tests. Some of these students may currently be receiving only native language (L1) literacy instruction, if allowable/available, in an effort to build literacy skills in L1 that will transfer to learning to read in English. Their comprehension is limited to

understanding simple isolated high frequency words and possibly short phrases. These beginning students are only able to comprehend very basic grammatical structures such as subject-verb-object sentences. They have extremely limited knowledge of how words, morphemes, and word order convey meaning in English.

English learners who fall somewhere between the low and medium levels of English proficiency best fit into the low proficiency group overall, given their characteristics, and thus also deserve attention here. The low/medium group includes those learners who can comprehend some short simple written phrases and sentences in familiar/social contexts that contain high frequency terms. However, they are not proficient at comprehending academic language or complex sentence structures. Also, their very limited vocabulary range necessitates frequent repetition and rephrasing for these learners' understanding; they frequently use contextual and visual cues to derive meaning and also rely largely on guessing. Thus, referring to and thinking about these students as "intermediate" may be misleading, as they typically require more extensive support than this label might imply.

General Accommodations Considerations for Low ELP Group

Given that English learners with low ELP have pre-functional to very limited English proficiency, the most salient and beneficial accommodations for the majority of students in this group are those that incorporate significant home language supports. As a result, native language or dual test forms are generally preferable to traditional forms in English, yet many students in this group may not yet have grade-level literacy in their L1, particularly learners in the elementary grades. For this reason, the addition of oral home language administration of the test will be an important and likely necessary accommodation to consider for students in this group. While access-based or "sheltered English" forms may be slightly preferable to traditional English test forms, even these modified forms are not likely to provide adequate access to testing for ELLs with low proficiency. In the event that no native language forms and/or access-based forms are available, oral administration in the L1 becomes even more essential; several states (e.g., Michigan) are utilizing this option of including an oral administration in L1 for subject areas other than reading for ELLs as a more feasible and cost-effective approach than translated test forms. A number of states that do not allow native-language accommodations (or instruction) do allow the translation or restatement of just the instructions, but no other parts of the test, for students. While admirable in its goal to make sure students understand the task demands, for ELLs with low English proficiency this practice is unlikely to be helpful since students who are unable to read the test questions in English (or in L1) will not be any better able to access the test and demonstrate their knowledge and skills than they would be without the instructions translated.

Tools to consider for English learners with low English language proficiency include picture/word dictionaries and possibly bilingual dictionaries, yet both of these choices should be predicated on the students' literacy levels in both the home language and English (and in reality are not likely to be as helpful for students in this ELP group as some of the other more comprehensive native language supports, e.g., oral administration of the test in L1). Manipulatives and other problem-solving tools, if applicable, are likely to be beneficial for some students in this group, particularly for items requiring constructed-response (CR). Computers offer many promising options for CR response modes, including drag-and-drop and model building features, but these types of test forms are not widely available at this time. In terms of response options, again, home language supports are ideal if not essential for students in this group: oral L1 response will be the most appropriate approach for ELLs with low or emergent literacy and writing skills in the home language and/or English. For students who possess adequate writing proficiency in the L1, providing the opportunity for ELLs to respond in writing in the native language and/or in code-switching is a viable option.

Additional administration accommodations to consider adding to the primary accommodations package include the more generic—yet no less important when thoughtfully combined with the other accommodations—practices of providing additional time, frequent breaks, and small group/individual administration (often with a familiar teacher to increase comfort level and/or to provide native language support). As Kopriva (2008) notes, "several authors (including Rivera & Stansfield, 2001; Duncan et al., 2005) emphasize that native language and dual language forms, as well as use of language aids (particularly if they are written), takes more time than if students are proficient in English and using English forms (p). She cites the example of test forms in Spanish, which "are approximately one-third longer than their counterparts in English" and goes on to state that "knowledge about these components in ELLs' home language reveals details about their level of proficiency which offers relevant information about these students' specific test-taking needs (LaCelle-Peterson & Rivera, 1994; Durán, 1997; Butler et al., 2001) (Kopriva 2008 p.). Kopriva (cite)has also recommended an additional administration accommodation involving the use of a specially trained "language liaison" to assist individual students for whom it is appropriate in reading/translating test forms and perhaps to provide explanation of particular non-target (construct-irrelevant) words, cultural contexts, and other factors that might not be comprehensible to these students, though she notes that effective training and a high degree of oversight to ensure test security are vital if such testing aides are used.

There is usually an interaction between administration and other accommodations (e.g., forms) and tools. For example, students using native language forms or dual forms and/or using bilingual dictionaries or other tools will almost certainly need additional time. However, additional time in isolation (and similarly, accommodations such as frequent breaks, small group, etc) is not likely to increase access to testing for ELLs, particularly those who are newly enrolled in U.S. schools and have had no opportunity to learn English and/or to acquire literacy in English or the home language (if L1 forms are allowable, available, and applicable, given state regulations and local capacity). Similarly, if they have not had the opportunity to learn subject matter in the content areas in the home country or in the U.S., extra time is likely to be inconsequential if not ineffective. In fact, for these students, unless oral/written native language supports are a) appropriate, and b) available, extra time could very well increase students' anxiety and/or fatigue, which is certainly not going to lead to more accurate test results.

Unfortunately, in many states, the only accommodations currently available for ELL students with the most nascent English proficiency include those that are highly unlikely to increase access to testing or validity of test inferences when administered. Several new tests and tools currently in research and development phases (ONPAR, etc) but not yet widely available are exploring the ways in which the dynamic capabilities of a computer-based testing platform can be especially beneficial to this group of learners with very beginning proficiency in English. Additional pretest support in the form of providing some orientation to the purposes and types of testing, not just to students but also to their parents, may be especially important and beneficial, given that students with the lowest levels of English proficiency are disproportionately represented by those who have not spent as much (or as consistent) time in the U.S. or in U.S. schools as their peers. For this reason, educators should pay special attention to the degree of and interplay between level of time in U.S., degree of cultural proximity, and U.S. schooling experiences for this group.

The low/medium group includes those learners who can comprehend some short simple written phrases and sentences in familiar/social contexts that contain high frequency terms. However, they are not proficient at comprehending academic language or complex sentence structures. Also, their very

limited vocabulary range necessitates frequent repetition and rephrasing for these learners' understanding; they frequently use contextual and visual cues to derive meaning and also rely largely on guessing. Thus, referring to and thinking about these students as "intermediate" may be misleading, as they typically require more extensive support than this label might imply.

Distinguishing Features of ELLs with *Medium* English Language Proficiency

Learners with medium ELP level understand non-academic English without much difficulty. However, they only understand some academic language and are just entering the initial stage of learning to comprehend complex sentence structures. They understand frequently used regular verbs and simple tenses but may be confused by irregular verb formation and non-simple tenses. Further, medium-level learners often have significant gaps in their range and use of vocabulary. Students in this group might also include those who have emergent or beginning literacy in English but are not yet reading at grade-level in English. In comparison with students in low or low/medium ELP groups, English learners with medium proficiency have slightly higher proficiency in academic English, syntax, and vocabulary. They are also a bit further along in their progress toward literacy in English, but they are still below grade level in this area.

General Accommodations Considerations for Medium ELP Group

Since learners with "medium" levels of English proficiency, as described above, have a fairly good grasp on non-academic English (i.e., Basic Interpersonal Communication Skills or BICS) but are still learning academic English, more complex sentence structures, grammatical features, and vocabulary, 'true' access-based or sheltered English forms are likely an appropriate accommodation to consider, particularly for students who are instructed solely in English and/or possess below grade level (or lower) literacy in the home language. Care should be taken, however, to ensure that available access-based forms include adequate language simplification with attention to the above linguistic features as well as appropriate compensatory visual support. While many states should be commended for having moved forward by providing more accessible or "universally designed" (UD) test forms, in most cases these types of UD forms, often standard forms that have been reviewed and revised by state-level or test publisher-employed specialists in ESOL and students with disabilities, are intended to be more accessible to the testing population as a whole, and do not contain sufficient changes that are significant enough to be beneficial for the target group of ELLs with medium ELP. Even true "plain language" or sheltered English forms, however, are not adequate in and of themselves. According to Kopriva (2008), she and Abedi (cite) "have repeatedly suggested that appropriate plain language and plain format forms are necessary but not sufficient. Proper language, tools and associated administration and response accommodations are also needed" (p.).

This is particularly true for students, often referred to as low intermediate or intermediate, whose proficiency in English falls somewhere between low and medium proficiency levels. For this reason (and especially for these "borderline" students), native language supports are often not only appropriate and valid but also essential for this group, depending on individual students' literacy levels in English and/or L1. For example, in the case of students in this group who are grade-level literate in the L1 and/or have received/are receiving instruction in the home language, native or dual language forms are likely the best option. Many students in this ELP group will likely derive benefits from oral administration in English or home language, as applicable, as long as they have aural proficiency levels, or as a supplement to access-based or native language forms accommodations. In the event that no sheltered forms or L1 forms exist or are allowed, as is often the case, oral administration options necessarily gain more primacy in the accommodations package.

Test-specific or general bilingual dictionaries may be helpful to this group to the extent they shore up some of the gaps in vocabulary that these students tend to have. However, such dictionaries are not likely to be adequate in the absence of sheltered English forms and/or native language forms or administration supports, given that students in this group are a) still learning to read in English, and b) still possess a beginning grasp of more complex grammar (e.g., prepositions and phrasal verbs, etc.) and sentence structure. Even if dictionaries may help them identify the meanings of some relevant words on the test form, dictionaries cannot supply the meaning-making and interpretation needed to understand the complex grammatical and discourse features that link these words together and synthesize essential information about what is required of the student in terms of comprehending what is being asked, figuring out an answer, and/or showing his or her knowledge and skills.

Since literacy in English for this group is still emergent and writing is so highly linked with literacy, it follows that response accommodations to be considered for students with low/medium and medium ELP should include access to native language supports such as written native language or code-switching or possibly oral response in English/L1. Demonstrated or modeled response is another viable option. Secondary administration accommodations, such as small group and extra time or additional breaks, may be necessary for students in this group relative to the combination of accommodations in their packages: again, dual or native language forms and the introduction of tools such as dictionaries will likely lengthen the amount of time needed to complete tests due to increased cognitive processing load (not to mention the additional load ELLs encounter regardless of the language of accommodations and/or test form as a consequence of their bilingualism and the fact that their brains are simultaneously processing linguistic information in both language systems). Finally, as with students in the low group, teachers should make an effort to address other test access issues that might arise for these students based on their level of cultural proximity, time in the U.S., and/or amount of experience with U.S. testing. In many cases, targeted pretest support providing orientation to testing and/or additional topics as applicable is merited; such support, while not necessarily in the form of a testing accommodation per se, nevertheless provides valuable assistance to ELLs depending on their individual background characteristics, needs, and strengths.

Distinguishing Features of ELLs with *High* English Language Proficiency

Learners with high ELP are able to comprehend and employ structures, vocabulary, and overall organizational patterns that approximate the speaking and listening of many native speakers of English in the classroom. These learners are usually rather comfortable with both familiar and unfamiliar topics. However, they may still require teacher support to achieve and maintain the pace of grade level reading and writing activities. Also, the learner with high ELP may lack grade level vocabulary terms that are specialized or infrequent. Within schools, this group of English learners typically includes those who are close to having grade-level reading and academic language proficiency in English. These ELLs are likely to meet state-level criteria for exiting ESOL support services by the close of the school year.

Unlike high students who are grade-level competitive by the end of the academic year, students who are higher than medium but not quite high often benefit from somewhat more intensive intervention and support. This medium/high group includes learners who understand non-academic English very well and academic English moderately well, but noticeably less well than native speakers of English. Specifically, they are often missing cultural language competences such as comprehending idioms, colloquialisms, and phrases and structures that require students to make inferences. These learners are able to understand most below grade level texts in content areas if care has been taken to remove problematic

language components such as those identified above. However, they are not entirely proficient at understanding complex language structures.

General Accommodations Considerations for High ELP Group

For those high students who are likely to be exited from ESOL services within the academic school year, the most salient accommodations are likely to be those that support emerging literacy and writing development. In particular, students may benefit from tools such as not only bilingual dictionaries but also test-specific English dictionaries (for tests other than English language arts). The goal is to bolster existing vocabulary and address any gaps that may persist; for example, these learners may still lack some specialized grade-level vocabulary and/or other academic English discourse features. Those at the top end of the ELP proficiency spectrum within the high ELP group can typically understand and construct language at a level that approximates native speakers—for this reason, this group of learners is sometimes referred to as "bridging" (WIDA) or grade-level competitive (Kopriva).

For the ELLs in the medium/high group (or the lower end of the high continuum), however, cultural competences, idiomatic linguistic structures, and structures requiring inferences may still present difficulty. Similarly, while bridging students are often comfortable with both familiar and unfamiliar topics, medium/high students, by contrast, are likely to have more success with more high-frequency topics, contexts, and vocabulary. Since these students can read most below grade-level texts provided that problematic linguistic features have been mitigated (e.g., sheltered English texts utilizing simple past, present, and future tenses with mostly simplified sentence structure and syntax), access-based or sheltered English forms may be an appropriate accommodation to consider. Likewise, judicious use of native language supports (especially dual forms and response options) may also be appropriate for some of the students on the medium/high borderline, depending upon their instructional experiences and literacy in L1. According to Kopriva (2008), "if and as they have some degree of written native language proficiency, written aid in L1 can supplement tests in English for those in English-speaking content classrooms. This because they can get clues about English item meaning from L1 cognates and other linguistic components (August, 2004)" (cite)." Kopriva then cites several other researchers (Solano-Flores and others (2002)) who maintain that "even when ELL students are rather literate in English…, "native language aids can continue to be helpful" (cite page of book). Thus, educators are encouraged to remain open to offering native language support—as appropriate based on literacy and proficiency in L1 and language of instruction—if it is available, even if and when students have relatively high proficiency in English.

Often, the "knee-jerk" response is to remove L1 supports from testing, if they are even available/legal, as soon as possible. In many states, English-only ideologies and/or policies may exacerbate the tendency to withhold or withdraw native language support prematurely. Further, the best-practice conventional wisdom is that students who are not or are no longer receiving instruction in the L1 should not get any native language testing supports. While in general this point of view has merit—ideally students should practice all accommodations in the classroom in a repetitive and ongoing manner if they are to be used validly and effectively in testing situations—in reality, the complex nature of second language acquisition, as discussed in the research literature (cite), suggests that cognitive differences in how bilinguals process information as well as asynchronous or asymmetrical development of English proficiency across not only domains (reading, writing, listening, speaking) but also content areas (math, science, etc.) may render bilingual and/or L1 supports appropriate and beneficial for even the highest level ELP students under certain circumstances. The feasibility and test validity issues associated with translated test forms and even oral administrations in L1 are well known, yet educators are encouraged to remain open-minded and let their own first-hand knowledge of individual students, along with

research on what works in terms of accommodations and how bilingual students read, guide their decisions about how best to meet the testing needs of their students.

It may be helpful to frame the decision-making around the notion that the goals of testing and instruction are not the same. If the goal of testing is to provide the most true and meaningful measure of what students know and can do in regard to academic content standards (as opposed to measure their progress toward learning English), then the objective should be to provide students with the most effective and useful supports that allow them to demonstrate their knowledge and skills with the least interference from any barriers imposed by their developing proficiency in English. This is because English learners do possess content knowledge and skills, despite the challenges they often face, and they deserve the chance to demonstrate them under circumstances that provide a more valid measure of what they know—because after all, if test results are to be useful, we must have faith in their accuracy and validity so that students' actual progress toward learning state academic standards and other learning objectives is measured. Allowing a student to use limited, highly targeted native language supports on a large-scale state test does not constitute a) providing an unfair advantage or b) keeping them isolated from their grade-level, native language peers in classes that deprive them of the chance to learn academic content at the same level of rigor—two arguments that represent the opposite spectrums of opposition to L1 instruction and testing. It may be helpful for educators and policy-makers to think carefully about the goals and purposes of large-scale testing in order to make thoughtful, objective decisions about accommodations that stem from thinking about student needs and strengths rather than policy, ideology, and/or expediency.

Selected Examples: Student Cases

The above discussion summarized some of the main salient characteristics of English learners at various stages of English proficiency. The next step in thoughtful and rigorous accommodations decision-making is to begin to distinguish accommodations needs among ELLs within the same ELP groups. As many researchers have noted, the English learner population in the United States is a highly diverse and heterogeneous group. Thus, if educators hope to provide the most valid and effective accommodations packages tailored to individual English learners, they must continue analyzing students' unique background characteristics and needs even after they are confident about their classification into the appropriate proficiency group. In order to differentially select accommodations for individual English learners within the same English language proficiency groups, it is necessary to take into account not just the critical features of their level of English proficiency as defined above, but also to carefully consider the additional student background factors previously discussed in this paper: time in U.S., culture, and level of experience with U.S. testing procedures.

The following discussion presents several examples of students within the same English proficiency group who nonetheless differ in important ways relative to accommodations decision-making. Considering "cases" or profiles of students with low to medium/high ELP will allow us to further unpack the complex (yet also surprisingly common-sense—e.g., "don't give a student an oral administration in English if her listening proficiency in English is low"— when broken into its constituent parts) decision-making that needs to take place in order to ensure ELLs access to testing. The following first two cases illustrate how crucial it is to think about varying levels of proficiency in both English and the L1, and further, to consider how proficiency in both languages may vary across the domains of reading, writing, listening and speaking. The next two examples will build on this theme while also demonstrating the importance of cultural proximity and experience with U.S. testing procedures.

Student Cases: Low English Proficiency

Consider the case of a fourth grader with a native language of Spanish who has just arrived in the United States two months into the academic school year due to the seasonal nature of the agricultural work done by her parents. This student has pre-functional English language proficiency in reading, writing, speaking, and listening. The student is well below grade level in L1 reading and writing proficiency due to sporadic schooling in the home country, given the cyclical nature of her parents' work in both Mexico and the U.S. Spanish is spoken exclusively in the home. Due to her recent enrollment in school, this student has obviously not had experience with standardized testing in the U.S. Because the student has pre-functional English proficiency, native language supports are essential if she is to access content-area assessments in English. However, assigning this student a Spanish or dual language form simply on the basis of her native language would not likely be beneficial given that her literacy level in the L1 is well below grade level; she is not likely to be able to fluently read test forms in Spanish OR in English. Rather, an oral administration of the test in Spanish is a more appropriate choice since her L1 listening proficiency in Spanish is likely much stronger because Spanish is utilized in the home. Note that even with an oral administration in Spanish, however, this student may still struggle to access the academic language in an L1 content test since she has not had consistent schooling or adequate opportunity to learn in her home country. Nevertheless, the L1 administration is still likely the best option for accommodating this student. Similarly, response options should ideally include the opportunity for the student to orally respond in the L1 to a scribe or language liaison, as defined by Kopriva (cite). Tools options for a student with this background may be limited in that bilingual and/or monolingual dictionaries will not be particularly beneficial since the student has well below grade-level literacy in English and Spanish, though manipulatives may be useful if applicable. Since the student has had inconsistent schooling in the home country and is newly enrolled in U.S. schools, cultural proximity and experience with U.S. testing are quite low. She (and likely her parents) would benefit from pretest support in the form of orientation to U.S. testing purposes and types via a "family assessment night" and/or conferences; the student will also benefit from one-on-one pretest classroom support from her teacher, including ongoing and repeated opportunities to practice typical U.S. large-scale test question types under conditions that should approximate the accommodations package she will receive on the state test.

Another student who has been identified as having low, beginner, and/or "Level 1" English proficiency may share many broad distinguishing features with the student cited above, but the ideal recommended accommodations package indicated for this student may look quite different. For example, let us assume that there is another student in the above student's fourth-grade class, also a native speaker of Spanish who has been enrolled in the U.S. less than one academic year. This student has pre-functional English language proficiency in reading and writing and beginner speaking and listening—in other words, he is similar (though just slightly higher in ELP listening and speaking) to the student described above. However, unlike the first student, this student has above grade level L1 reading proficiency, having had access to consistent and high-quality educational opportunities in his home country prior to coming to the U.S. with his family. In this instance, then, a Spanish or dual language form would be a viable accommodation to consider for this student. Similarly, given his proficiency in Spanish, a bilingual Spanish-English dictionary might be useful for this student if he is receiving a dual form. In the event a dictionary tool is used, the student may need additional time or breaks given the longer length of Spanish forms as compared to English and the processing demands of using a dictionary. This second student may also benefit from an oral administration of the test in the L1 (though it would not be critical, unlike for the first student, given his high literacy in L1), either in addition to the dual or native language form, or as a 'second-choice' substitute for the L1/dual form, if it is not available. In terms of response accommodations, allowing the second student to respond in writing in L1 would be advisable; oral responses in L1 would likely be an appropriate secondary choice, assuming the student has at least

grade level oral proficiency in the home language (again, he may possess higher social than academic Spanish, but in this case the academic Spanish discourse is likely to be more developed due to his schooling in the home country).

Unfortunately, if no native language supports are allowable or available in the contexts in which these students attend school, standard or even sheltered English forms (which are typically combined with oral administration in English, small group, extra time, and dictionaries in the absence of other more relevant and effective native language supports), are likely to be ineffective. This does both students a disservice but is uniquely painful in the case of the student who has had opportunity to learn rigorous academic content in his native country: while he very likely does have knowledge and skills in subject areas, without native language supports he will be utterly incapable of demonstrating this, thereby providing an inaccurate picture of his learning and thinking to his teacher, policymakers and other stakeholders—with all the possible short and long-term negative repercussions that can entail.

Student Cases: Medium English Proficiency

For the sake of discussion, let us consider the case of a sixth-grader with a native language of Japanese. The student has attended school regularly in the United States for one academic year, but attended a Japanese school with a strong (American) English language program with U.S. teachers who employed a communicative (not solely literacy and/or drill-based approach) and many of the same resources and instructional and testing procedures as the U.S. The student has high intermediate English language proficiency in reading, writing, and listening and is at grade level in L1 reading proficiency. The student has had experience with all of the relevant standardized testing procedures in the home country except answering open ended questions requiring the *student's* ideas; he also struggles when U.S. assignments and questions that require him to take a stand/make an argument on a particular issue, particularly if his opinion contradicts that of his teacher. In other words, he is still adjusting to cultural dissonance related to his experience as a member of a (home) culture that values the communal good over the individual and in which elders and superiors are highly respected--which at times conflicts with the individualistic culture and value system of the U.S., and by association, the U.S. educational system.

Contrast this student with his classmate, an sixth-grade girl with a native language of Cantonese who has attended school regularly in the United States for more than three academic years but who lives in a tight-knit Cantonese community that, while in the U.S., is culturally—and linguistically--still almost identical to the home country. This student has ELP levels similar to those of her classmate, with high intermediate English language proficiency in reading and writing and slightly lower levels of speaking and listening. She is below grade level in L1 reading proficiency, having not had the opportunity for formal schooling in her L1 since she came to the U.S. at a young age from a remote and poor farming region in China with few schools physically accessible to children (who typically assisted their parents in cultivating land rather than attending school regularly due the distance of the school—and lack of transportation--from their home). The student has had experience with all of the relevant standardized testing procedures given her three years in the U.S. Because she has been in the U.S. for several years, she has had experience with the testing procedures and types of questions on U.S. tests, yet she still has trouble "getting" certain passages or contexts on tests that automatically assume common U.S. background and cultural information, given the community in which she currently lives.

While both of these students have similar levels of English proficiency, several important differences merit consideration as they relate to accommodations decision-making. First, the Japanese student, while having spent less time in the U.S. (one academic year) compared to the Cantonese-speaking student (over three years), actually has had more culturally similar schooling experiences than the

Cantonese student. This is due to the nature of the schooling he received in his home country, which had many Western features, as opposed to the Cantonese student, who has had the majority of (and more consistent) schooling in the U.S. but who lives in an American community that is culturally Cantonese. Both students may still benefit from some accommodations and pretest support that address cultural and testing experience issues, but for different reasons, which will yield accommodations packages that might not look the same. For instance, despite the fact that the Japanese student has been in the U.S. substantially less time than the Cantonese student, he may require less support that addresses cultural proximity issues as they relate to texts and other contexts typically found in U.S. test items than his classmate, who may have more difficulty with assumptions of construct-irrelevant cultural background knowledge due to the insular nature of her home community; both may have difficulty managing assumptions of common value systems—e.g., communal vs. individual values. The Japanese student will probably benefit from some pretest support in the form of being reassured by his teacher that it is not only appropriate but necessary for him to be critical and/or take stands on issues in order to be successful in the U.S. educational and testing system.

In conclusion, it is important for educators to thoroughly think through the issue of how much time students have been in the U.S. as well as their home and school cultures along with other factors. As this example illustrated, time in U.S. as a student factor is not as meaningful if it is not considered in the context of the student's other life and schooling circumstances, and thus assumptions should not be made about which accommodations to give students solely on the basis of how much time they have spent in the U.S. or what type of cultural background they come from. Rather, educators must think about the whole person when making testing accommodations decisions—just as they do so naturally every day in their classrooms. With a bit of training and support, translating this intuitive grasp of student needs and strengths as it relates to instruction to testing accommodations is not such a stretch but rather a natural extension of what teachers do best of all.

Other Considerations: Capacity, e.g., Local L1 Instructional and Testing Capacity; Local and State Constraints on Allowable/Available

What is STELLA/Relevance of STELLA

STELLA is an online decision-making tool designed to assist educators in selecting appropriate testing accommodations for English learners. The STELLA system includes decision-making algorithms for determining the specific accommodations individual student should receive on large-scale tests for accountability purposes. In order to recommend accommodations based on students' particular background characteristics and needs, STELLA collects data from three sources: school records, teachers' input and parents' input. This information includes students' familiarity with standardized academic assessments, formal and informal English language proficiency (ELP) data, primary language (L1) proficiency data, similarity of schooling system in native country to U.S. schools, and time attending school in the United States. In the STELLA system, gathering data from multiple sources is a way of improving the accuracy of the student profile and helping ensure that the right set of accommodations is matched to that particular student. The Records Form collects information from the student's file at the school, such as scores on English language proficiency and content tests. The Parent/Guardian Interview Form guides an educator through an interview process to collect data from the student's parent or guardian. The Teacher Form collects information from the teacher(s) who works directly with the student in the classroom.

STELLA consolidates this information and creates a profile for each student. The system then selects one or more accommodations deemed most appropriate for the student from a set of approximately twenty promising research-based and practitioner-recommended testing options. The system output displays recommended state-allowed accommodations for individual students and grays out accommodations that are STELLA-recommended but not permitted by the state. These recommended research-based accommodations are intended to support and guide–not replace–the decision-making processes of educators involved in assigning accommodations to individual English learners on large-scale assessments.

Purpose of STELLA/How It Can Help

STELLA was created to provide a systematic, research-based method for determining appropriate testing accommodations for individual ELL students on large-scale assessments. The overall purpose of STELLA is to support consistent, research-based, and fair decision-making processes that benefit both students and the educators who instruct them. The STELLA collects and analyzes relevant student background information to match individual learners with valid accommodations for participation in state and district tests. These accommodations represent changes in the testing environment, procedures, or presentation of items or tests that do not alter the targeted content or rigor measured by the items, and thus yield valid scores that are comparable with non-accommodated scores. In this way, STELLA can help encourage the meaningful inclusion of ELL students in large-scale assessments for accountability purposes.

Growing research indicates that utilizing valid, fair, and effective accommodations leads to more accurate test scores that better reflect what ELLs know and can do in regard to state content standards–rather than the extent to which their emerging English language proficiency may impede their access to the test and/or their ability to demonstrate their content knowledge and skills. Having accurate results for these and all students is essential to ensure the accountability of the educational system, determine how to best meet students' educational needs, and track student progress over time.

Conclusions: How Learning about STELLA Help Educators with Accommodations Decision-making?

At this point it should be clear that options in the ways test materials are presented and administered is very important for students who come from diverse cultural backgrounds and have English literacy and language challenges. However, no matter how adroitly and carefully these options are selected and created, if they don't get to the correct students who need them, they are not useful. Emerging research suggests that use of improper accommodations may result in scores that are not substantially more valid for these students than if they received no accommodations. This is very troubling for many reasons but particularly two. First, individual performance on large scale academic tests and scores may misrepresent the skills of students who do not receive proper accommodations. This can have profound consequences for future learning opportunities for these students, including improper placement or other student level decisions, as well as inappropriately impacting educational agency level decisions and possibly resource allocations. Second, improper accommodations may help explain why the findings from accommodations research have, to date, been so mixed. It may be suggested that only when researchers employ rigorous means of matching students to accommodations they are empirically testing will their work provide clearer results about which options are useful and which are not. Currently, many of us are left with the vague notion of which accommodations are helpful but we're not sure how or when. The real question may well be "Useful for whom?" It is argued here that this is the question researchers should be investigating.

Finally, some researchers envision accommodation assignment approaches that are integrated into computerized content area testing systems. Kopriva, Carr and Cho (2006) discuss future testing systems that are sensitive to a number of nuances about the ELL students and alter accommodations by item. These systems would not only allow for matching algorithms to operate in them, such as that produced by STELLA, but they would include algorithms that address item characteristic data which are more finely tuned to student needs. For instance, Solano-Flores, Trumbull, and Nelson-Barber (2002) suggest that ELL students differentially use native language or English items when presented with dual language forms. While others (e.g. Duncan et al., 2005) suggest students use primarily one form or another, Solano-Flores suggests that his more fine-grain research point to small preferences depending on a number of linguistic factors.